Characterizing and computing real pseudospectral abscissa for nonlinear eigenvalue problems

Francesco Borgioli
Wim Michiels
Nicola Guglielmi

Report TW 668, August 2016

KU Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Characterizing and computing real pseudospectral abscissa for nonlinear eigenvalue problems

Francesco Borgioli
Wim Michiels
Nicola Guglielmi

Report TW 668, August 2016

Department of Computer Science, KU Leuven

Abstract

We present a new iterative algorithm to compute the pseudospectral abscissa for a class of nonlinear eigenvalue problems, which includes the delay and the polynomial eigenvalue problem. Two methods are developed that exploit real valued matrix perturbations, for the Frobenius and the spectral norm. In both cases, it is proved that the pseudospectral abscissa can be obtained when restricting to at most rank-2 perturbations and that the critical perturbations on the coefficient share their column and row spaces. This is a surprising property, giving that the combined additive perturbation on the characteristic matrix is highly structured. The proposed iterative algorithms can be interpreted as the discretization of a gradient ow, where properties of critical perturbations are fully exploited. Even though the main contributions concern the nonlinear eigenvalue problem, the paper contributes to the special case of the standard eigenvalue problem in the following way. For the Frobenius norm perturbations we provide differential equations generating a path in the space of perturbations of rank smaller or equal to two (instead of rank exactly two in the present literature), allowing us to treat the case of complex and real rightmost eigenvalues simultaneously. For the spectral norm case we derive explicit optimality conditions in terms of the compact singular value decomposition. Finally we provide two extensions for the Frobenius norm pseudospectrum. The first one consists of using a combined measure on the perturbations, allowing us to assess from the critical perturbations which coefficient matrices the pseudospectral abscissa is most sensitive to. The second extension consists of considering structured perturbations on the coefficient matrices. In this way our algorithm exploits structure at three levels: the structure of the nonlinear eigenvalue problem (instead of studying unstructured pseudospectra of some linearization of the eigenvalue problem), the property that perturbations are assumed to be real valued, and additional structure on the perturbations of individual matrices. For this case we show that the pseudospectral abscissa cannot always be reached when restricting to perturbations of maximum size.

Keywords : nonlinear eigenvalue problem, pseudospectral abscissa, low-rank perturbations.
MSC : 47J10, 65F15.

Characterizing and computing real pseudospectral abscissa for nonlinear eigenvalue problems

Francesco Borgioli* Wim Michiels* Nicola Guglielmi†

Abstract

We present a new iterative algorithm to compute the pseudospectral abscissa for a class of nonlinear eigenvalue problems, which includes the delay and the polynomial eigenvalue problem. Two methods are developed that exploit real valued matrix perturbations, for the Frobenius and the spectral norm. In both cases, it is proved that the pseudospectral abscissa can be obtained when restricting to at most rank-2 perturbations and that the critical perturbations on the coefficient share their column and row spaces. This is a surprising property, giving that the combined additive perturbation on the characteristic matrix is highly structured. The proposed iterative algorithms can be interpreted as the discretization of a gradient flow, where properties of critical perturbations are fully exploited. Even though the main contributions concern the nonlinear eigenvalue problem, the paper contributes to the special case of the standard eigenvalue problem in the following way. For the Frobenius norm perturbations we provide differential equations generating a path in the space of perturbations of rank smaller or equal to two (instead of rank exactly two in the present literature), allowing us to treat the case of complex and real rightmost eigenvalues simultaneously. For the spectral norm case we derive explicit optimality conditions in terms of the compact singular value decomposition.

Finally we provide two extensions for the Frobenius norm pseudospectrum. The first one consists of using a combined measure on the perturbations, allowing us to assess from the critical perturbations which coefficient matrices the pseudospectral abscissa is most sensitive to. The second extension consists of considering structured perturbations on the coefficient matrices. In this way our algorithm exploits structure at three levels: the structure of the nonlinear eigenvalue problem (instead of studying unstructured pseudospectra of some linearization of the eigenvalue problem), the property that perturbations are assumed to be real valued, and additional structure on the perturbations of individual matrices. For this case we show that the pseudospectral abscissa cannot always be reached when restricting to perturbations of maximum size.

1 Introduction

Characterizing and computing pseudospectra are well established tools for analyzing the sensitivity of eigenvalues of a system [20]. Pseudospectra are sets in the complex plane to which the eigenvalues can be moved when the system is subject to perturbations with a specified upper bound. The pseudospectral abscissa constitutes a bound on the asymptotic growth rate of the solutions of the perturbed system, which is uniform over all possible perturbations under consideration and, therefore, it allows to assess the stability robustness [5]. The pseudospectral abscissa is hence closely related to the distance to instability [1, 2].

*Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, 3001 Heverlee, Belgium ({firstname.lastname}@cs.kuleuven.be).

†Dipartimento di Ingegneria, Scienze Informatiche e Matematica and Gran Sasso Science Institute, Università dell'Aquila, via Vetoio (Coppito), I-67010 L'Aquila, Italy (guglielm@univaq.it).

In this article we analyze a wide class of perturbed nonlinear eigenvalue problems that includes the polynomial eigenvalue problem and the delay eigenvalue problem, more precisely

$$\left(\sum_{i=1}^{m}(A_i+\delta A_i)p_i(\lambda)\right)y=0,\ \lambda\in\mathbb{C},\ y\in\mathbb{C}^n, \tag{1.1}$$

where $A_i \in \mathbb{R}^{n\times n}$ with $n \geq 2$, and function p_i is assumed entire, having a Taylor expansion around zero with real coefficients only, for $i = 1, \ldots, m$. In the characterization of pseudospectra, our aim is to maximally exploit the structure of the eigenvalue problem and of the perturbations $\delta A_i,\ i = 1, \ldots, m$, in the analysis. A traditional approach to handle nonlinearity in the frequency domain consists of applying a so-called linearization of the nonlinear eigenvalue problem, possibly in combination with a (usually polynomial or rational) approximation, which results in a standard eigenvalue problem, followed by the pseudospectral abscissa computation using standard techniques. However, this approach has two drawbacks. Firstly, it does not necessarily simplify the problem. Indeed, linearizations have been widely studied for polynomial eigenvalue problems, but few results are present regarding generic nonlinear eigenvalue problems [7], and the linearization may result in an infinite-dimensional system, while the combination with an approximation introduces an approximation error which may be hard to characterize. For example, a linearization of the delay eigenvalue problem gives rise to an infinite-dimensional eigenvalue problem of a differentiation operator with non-local boundary conditions [15]. Secondly, the analysis of unstructured pseudospectra of the linearization does not respect the structure of equation (1.1), leading to, among others, an over-bounding of the actual uncertainty and conservative robust stability criteria.

In [6] and [11] two different methods are proposed for the pseudospectral abscissa computation of (1.1) that fully respect the structure of equation, but restrict the analysis to complex perturbations. These methods, although different in nature (based on constructing low rank perturbations converging to the critical perturbations, thereby generalizing the algorithm of [5] for the standard eigenvalue problem, and based on eigenvalue optimization techniques for Hermitian eigenvalue problems as in [10]), strongly rely on the singular value characterization of pseudospectra, developed in [13], as a generalization of [19] for polynomial eigenvalue problems. In this paper, we make the leap from complex valued to real valued perturbations, which are more realistic from an application point of view. To illustrate that the generalization is far from trivial, we note that the combined additive perturbation of the characteristic matrix is given by

$$\Delta_{\text{ad}} := \sum_{i=1}^{m}\delta A_i\ p_i(\lambda). \tag{1.2}$$

If the perturbations δA_i are allowed to be complex valued, then the functions p_i can be "absorbed" in the matrices A_i. This observation, which is at the basis of the singular value characterization in [13], makes that only information on the modulus of functions $p_i(\lambda)$ is exploited. On the contrary, if the perturbations are real valued (but some on the p_i are complex valued), the additive perturbation Δ_{ad} has a special structure. In the last section of the paper, we exploit the additional property that perturbations δA_i may be structured. The latter is particularly useful when the eigenvalue problem (1.1) stems from a system of (delay) differential algebraic equations, modeling complex interconnected systems [9, 12], where the structure on the perturbation may stem from physically relevant perturbations to one of the components or interconnections.

In this work we start by deriving optimality conditions for rightmost points of the pseudospectrum, where perturbations on the individual coefficient matrices are constrained using

both the Frobenius and the spectral norm. We demonstrate that for both cases, it is possible to restrict the sets of optimal perturbations that produce the rightmost points as sets of up to rank-two matrices. Moreover, we show that the optimal perturbations on the different matrices A_i share their column and row spaces. These surprising properties, giving the complicated structure of (1.2), allow us to extend gradient based optimization algorithms, exploiting low rank dynamics, developed in [3] and [4], for complex, respectively real pseudospectra of the standard eigenvalue problem. These algorithms rely on the construction of a steepest ascent differential equation in the space of perturbations, aimed at maximizing the real part of the rightmost eigenvalue. For the Frobenius norm case, we provide a differential equation in the space of perturbations of rank smaller or equal to two (in contrast to an iteration on the Riemannian manifold of matrices of rank exactly two, endowed with the "natural" geometry, in [4]), which allows us to treat real and complex rightmost eigenvalues simultaneously. The novel optimal conditions which we give for the spectral norm case are tightly connected with the fixed points of the iteration derived in [17] for computing the real pseudospectral abscissa of a matrix. The generalization of the iteration to the nonlinear eigenvalue problems will play a role in the initialization of our algorithm.

The paper is structured as follows: section 2 presents the problem and states the main assumptions. Sections 3 and 4 show the theoretical results that lead to the algorithm construction for the Frobenius norm and the spectral norm case, respectively. Section 5 describes the results obtained in several numerical experiments. Section 6 is devoted to two extensions of the algorithm for the Frobenius norm case, dealing with a different measure on the combined perturbation $(\delta A_1, \ldots, \delta A_m)$, and with structured perturbations on the coefficient matrices, respectively. Section 7 presents the conclusions.

2 Problem statement and main assumptions

In order to characterize pseudospectra of (1.1) we need to define a scalar measure on the set of perturbations. Introducing $\Delta = (\delta A_1, \ldots, \delta A_m)$ and weights $w_i \in \overline{\mathbb{R}}_0^+ = \mathbb{R}_+ \setminus \{0\} \cup \{\infty\}$, we define the following norm

$$\|\Delta\|_{\text{glob}} := \left\| \begin{bmatrix} w_1 \|\delta A_1\| \\ \vdots \\ w_m \|\delta A_m\| \end{bmatrix} \right\|_\infty, \tag{2.1}$$

where $\|\cdot\| = \|\cdot\|_2$ or $\|\cdot\| = \|\cdot\|_F$. Therefore, asking for an ϵ-bounded set of perturbation is equivalent to impose

$$\|\delta A_i\| \leq \frac{\epsilon}{w_i}, \ i = 1, \ldots, m.$$

By choosing some of the weights equal to infinity, we can impose that some coefficient matrices are not perturbed.

Given the above measure, we can now introduce the notion of real ϵ-pseudospectrum as a set

$$\Lambda_\epsilon := \bigcup_{\substack{\Delta \in \mathbb{R}^{n \times n \times m}, \\ \|\Delta\|_{\text{glob}} \leq \epsilon}} \left\{ \lambda \in \mathbb{C} : \det\left(\sum_{i=1}^{m} (A_i + \delta A_i) p_i(\lambda) \right) = 0 \right\}.$$

The corresponding pseudospectral abscissa is defined as

$$\alpha_\epsilon := \sup\{\Re(\lambda) : \lambda \in \Lambda_\epsilon\}. \tag{2.2}$$

Throughout the paper we make the following assumption on the shape of the ϵ-pseudospectra.

Assumption 2.1. *The set* $\Lambda_\epsilon \cap \{\lambda \in \mathbb{C} : \Re(\lambda) \geq r\}$ *is bounded for all* $r \in \mathbb{R}$.

Assumption 2.1 guarantees that a globally rightmost point of the pseudospectrum exists, hence, the supremum operator can be replaced by a maximum in (2.2). We refer to [14, Section 2] for an elaborate discussion on this assumption.

Secondly, an assumption is made that plays an important role in the characterization of the rank of critical perturbations that we will address in the next sections.

Assumption 2.2. *Let* λ_{RM} *be a rightmost eigenvalue of (1.1), and let* x, y *be its corresponding left and right eigenvectors respectively. If* $\Im(\lambda) \neq 0$*, then vectors* $\Re(x)$ *and* $\Im(x)$ *are independent, as well as vectors* $\Re(y)$ *and* $\Im(y)$*. Equivalently, matrices*

$$X = [\Re(x)\ \Im(x)],\ Y = [\Re(y)\ \Im(y)]$$

in $\mathbb{C}^{n\times 2}$ *have maximum rank.*

Even though this assumption always holds for the standard eigenvalue problem (recall that $n \geq 2$) and it sounds very natural, it is not satisfied for all nonlinear eigenvalue problems under consideration. As a counter example, consider the quadratic eigenvalue problem

$$\left(\lambda^2 I + \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\right) y = 0,$$

where rightmost complex eigenvalues admit a purely real eigenvector.

Finally, we make the following, technical assumption.

Assumption 2.3. *Let* λ_{RM} *be a rightmost point of the* ϵ*-pseudospectrum. Then we have* $p_i(\lambda_{\mathrm{RM}}) \neq 0$ *whenever* w_i *is finite, for* $i = 1, \ldots, m$.

In [14] this assumption is motivated by showing that the violation corresponds to a degenerate case (the conclusions of Proposition 3.3 in [14] carry over from complex to real pseudospectra since they share the intersection with the real axis). For the special case of the polynomial eigenvalue problem, a violation would imply that $\lambda_0 = 0$ and that this eigenvalue is invariant under perturbations.

3 The Frobenius norm case

3.1 Foundations of the algorithm

To construct our algorithm we make use of the following two lemmas, whose proof can be easily constructed from Lemma 2.7 in [18]. In this section and in the following ones we will refer to the Frobenius inner product using the symbol $\langle \cdot, \cdot \rangle$.

Lemma 3.1. *Let* $t \in \mathbb{R}_+$ *and consider the nonlinear eigenvalue problem*

$$\left(\sum_{i=1}^{m} \big(A_i + \delta A_i(t)\big) p_i(\lambda)\right) y = 0,$$

where the functions $t \to \delta A_i(t)$ *are assumed differentiable for* $i = 1, \ldots, m$*. Let* λ *be a simple eigenvalue of the system, and let* x *and* y *be its corresponding left and right eigenvectors, then we have*

$$\frac{d\lambda(t)}{dt} = -\frac{x^*\Big(\sum_{i=1}^{m} \big(\delta \dot{A}_i(t)\big) p_i(\lambda)\Big) y}{x^*\Big(\sum_{i=1}^{m} \big(A_i + \delta A_i(t)\big) p_i'(\lambda)\Big) y}. \tag{3.1}$$

Lemma 3.2. *Consider non-linear eigenvalue problem (1.1), where* $\delta A_i \in \mathbb{R}^{n\times n},\ i = 1,\dots,m$ *are matrices defined by* n^2 *independent variables such as*

$$\delta A_i = \begin{pmatrix} v_{1,1}^{(i)} & \dots & v_{1,n}^{(i)} \\ \vdots & \ddots & \vdots \\ v_{n,1}^{(i)} & \dots & v_{n,n}^{(i)} \end{pmatrix}.$$

Let λ *be a simple eigenvalue of the system, and let* x *and* y *be its corresponding left and right eigenvectors. Then we have*

$$\frac{\partial \lambda}{\partial v_{s,t}^{(k)}} = -\frac{\bar{x}_s p_k(\lambda) y_t}{x^*\Big(\sum_{i=1}^m (A_i + \delta A_i) p_i'(\lambda)\Big) y},\ k = 1,\dots,m,\ s,t = 1,\dots,n. \tag{3.2}$$

The next theorem represents the main theoretical result of this section, and the construction of the iterative method is based on it.

Theorem 3.3. *Let* λ_{RM} *be a globally rightmost point in the real* ϵ*-pseudospectrum and assume it is a simple eigenvalue for* ϵ*-bounded perturbations* $(\delta A_1,\dots,\delta A_m)$*. Let* x *and* y *be its corresponding left and right eigenvectors normalized such that*

$$x^*\left(\sum_{i=1}^m (A_i + \delta A_i) p_i'(\lambda_{\mathrm{RM}})\right) y > 0. \tag{3.3}$$

Defining

$$X = [\Re(x)\ \Im(x)],\ Y = [\Re(y)\ \Im(y)],\ \Gamma_i = \begin{bmatrix} \Re(p_i(\lambda_{\mathrm{RM}})) & -\Im(p_i(\lambda_{\mathrm{RM}})) \\ \Im(p_i(\lambda_{\mathrm{RM}})) & \Re(p_i(\lambda_{\mathrm{RM}})) \end{bmatrix},\ i = 1\dots m,$$

we have $X\Gamma_i Y^T \neq 0$ *and we can express the optimal perturbations as*

$$\delta A_i = -\frac{\epsilon}{w_i}\frac{X\Gamma_i Y^T}{\|X\Gamma_i Y^T\|_F}. \tag{3.4}$$

Proof. Let us consider the perturbed nonlinear eigenvalue problem

$$\left(\sum_{i=1}^m (A_i + \delta A_i) p_i(\lambda)\right),$$

and let us define every perturbation matrix $\delta A_i \in \mathbb{R}^{n\times n}$ as depending on n^2 independent variables

$$\delta A_i = \begin{pmatrix} v_{1,1}^{(i)} & \dots & v_{1,n}^{(i)} \\ \vdots & \ddots & \vdots \\ v_{n,1}^{(i)} & \dots & v_{n,n}^{(i)} \end{pmatrix},\ i = 1,\dots,m.$$

Using Lemma 3.1 and defining $\xi := x^*\Big(\sum_{i=1}^m (A_i + \delta A_i) p_i'(\lambda_{\mathrm{RM}})\Big) y$, we can state that

$$\frac{\partial \Re(\lambda_{\mathrm{RM}})}{\partial \delta A_i} := \left(\frac{d\Re(\lambda_{\mathrm{RM}})}{dv_{s,t}^{(i)}}\right)_{s,t=1,\dots,n} = -\frac{1}{\xi}\left(\bar{x}_s p_i(\lambda_{\mathrm{RM}}) y_t\right)_{s,t=1,\dots,n} = -\frac{1}{\xi} X\Gamma_i Y^T \neq 0,$$

where we employed Assumption 2.2. We want all our perturbations to be ϵ-bounded in Frobenius norm, thus we impose on each of them the following constraint,

$$g_i := \sum_{s,t=1,\dots,n} \left(v_{s,t}^{(i)}\right)^2 - \frac{\epsilon^2}{w_i^2} \leq 0,\ i = 1,\dots,m. \tag{3.5}$$

Since λ_{RM} is a globally rightmost point, the following optimality conditions are satisfied,

$$\frac{\partial \mathfrak{R}(\lambda_{\mathrm{RM}})}{\partial \delta A_i} - 2\mu_i \delta A_i = 0,\ i = 1,\dots,m,$$

where $\mu_i \geq 0,\ i = 1,\dots,m$ are the Lagrange multipliers corresponding to (3.5). Since $\frac{\partial \mathfrak{R}(\lambda_{\mathrm{RM}})}{\partial \delta A_i} \neq 0$, we must have $\mu_i > 0,\ i = 1,\dots,m$. We conclude that the critical set of perturbations must lie on the boundary of the ϵ-radius ball, and be a positive multiple of $\frac{\partial \mathfrak{R}(\lambda_{\mathrm{RM}})}{\partial \delta A_i}$.

Introducing weights $w_i \in \overline{\mathbb{R}}_0^+$ and imposing $\|\delta A_i\|_F = \dfrac{\epsilon}{w_i},\ i = 1,\dots,m$, we get the assertion of the theorem. □

The previous result characterizes the critical set of perturbations using the rightmost point of the pseudospectrum and its left and right eigenvector. This provides two fundamental results to build the algorithm. Firstly, the critical perturbations have maximum norm and are rank-two if the rightmost eigenvalue λ_{RM} is complex, while in case of a real λ_{RM} the critical δA_i have rank equal to one. This allows us to find the optimal perturbations by restricting to matrices defined on a manifold as follows:

$$\mathcal{S}_F := \{(\delta A_1,\dots,\delta A_m) \in \mathbb{R}^{n\times n\times m} : \mathrm{rank}(\delta A_i) \leq 2 \text{ and } w_i \|\delta A_i\|_F = \epsilon,\ i = 1,\dots,m\}. \tag{3.6}$$

Secondly, all the optimal perturbations δA_i share the same column space with the matrix X and the same row space with the matrix Y^T.

In order to reach the optimal perturbations, we construct a path along $\mathcal{S}_F$ where perturbation matrices are defined through the compact singular value decomposition

$$\delta A_i(t) = -\frac{\epsilon}{w_i} \widehat{U}_i(t) \widehat{\Sigma}_i(t) \widehat{V}_i(t)^T,\quad t \in \mathbb{R}_+,$$

where $\widehat{U}_i(t), \widehat{V}_i(t) \in \mathbb{R}^{n\times 2}$ have orthonormal columns and $\widehat{\Sigma}_i(t) \in \mathbb{R}^{2\times 2}$ is a diagonal matrix such that $\left\|\widehat{\Sigma}_i(t)\right\|_F = 1$. Let us observe that

$$\dot{\widehat{U}}_i(t)^T \widehat{U}_i(t) + \widehat{U}_i(t)^T \dot{\widehat{U}}_i(t) = 0,\ \forall\ t \geq 0,$$

thus $\widehat{U}_i(t)^T \dot{\widehat{U}}_i(t)$ is 2×2 skew-symmetric matrix, and the same result applies for $\widehat{V}_i(t)^T \dot{\widehat{V}}_i(t)$. To build our iterative method we will make use of the following result.

Proposition 3.4. *Let the function $\mathbb{R}_+ \ni t \to \widehat{U}(t) \in \mathbb{C}^{n\times 2}$ be continuous for all t and differentiable almost everywhere, such that the columns of $\widehat{U}(t)$ are orthogonal for all $t \in \mathbb{R}_+$. Then there exist a function $t \ni \mathbb{R}_+ \mapsto Q(t) \in \mathbb{C}^{2\times 2}$, with $Q(t)$ orthogonal for all t, and a function $\mathbb{R}_+ \ni t \mapsto U(t) \in \mathbb{C}^{n\times 2}$ such that $\widehat{U}(t) = U(t)Q(t)$ and $U(t)^T \dot{U}(t) = 0\ \forall\ t$.*

Proof. Whenever $\widehat{U}$ is differentiable we have, following from the orthogonality property, $\widehat{U}^T \dot{\widehat{U}}$ is skew-symmetric. Let $Q(t)$ be a solution to the following Cauchy problem

$$\begin{cases} \dot{Q}(t) = -Q(t)[\widehat{U}(t)^T \dot{\widehat{U}}(t)]^T, & \text{if } t > 0, \\ Q(0) = I_2. \end{cases}$$

With this definition $Q(t)$ is orthogonal, thus $Q(t)^T\dot{Q}(t) = -[\widehat{U}(t)^T\dot{\widehat{U}}(t)]^T$. Then we define $U(t) := \widehat{U}(t)Q(t)^T$, so that

$$U(t)^T\dot{U}(t) = Q(t)\widehat{U}(t)^T\dot{U}(t) = Q(t)\dot{Q}(t)^T + Q(t)\widehat{U}(t)^T\dot{\widehat{U}}(t)Q(t)^T =$$

$$= -Q(t)[\widehat{U}(t)^T\dot{\widehat{U}}(t)]Q(t)^T + Q(t)\widehat{U}(t)^T\dot{\widehat{U}}(t)Q(t)^T = 0.$$

□

Thanks to this proposition and to the fact that the Frobenius norm is invariant to orthogonal transformations, we can rewrite the previous decomposition as follows,

$$\delta A_i(t) = -\frac{\epsilon}{w_i}U_i(t)Q_i(t)V_i(t)^T,$$

with $U_i(t), V_i(t) \in \mathbb{R}^{n\times 2}$ s.t. $U_i(t)^T\dot{U}_i(t) = V_i(t)^T\dot{V}_i(t) = 0$ and $Q_i(t) \in \mathbb{R}^{2\times 2}$ s.t. $\|Q_i(t)\|_F = 1\ \forall\ t \geq 0,\ i = 1, \ldots, m$.

Let us observe that conditions

$$U_i(t)^T\dot{U}_i(t) = V_i(t)^T\dot{V}_i(t) = 0\ \forall\ t \geq 0,\ i = 1, \ldots, m$$

can be imposed, without loosing generality, by a parameterization

$$\begin{cases} \dot{U}_i(t) = \Big(I_n - U_i(t)U_i(t)^T\Big)R_i(t) \\ \dot{V}_i(t) = \Big(I_n - V_i(t)V_i(t)^T\Big)S_i(t), \end{cases} \tag{3.7}$$

for any $R_i(t), S_i(t) \in \mathbb{R}^{n\times 2}$.

Furthermore, the condition $\|Q_i(t)\|_F = 1\ \forall\ t \geq 0,\ i = 1, \ldots, m$ is equivalent to require $Q_i(t)$ to be orthogonal to $\dot{Q}_i(t)$ with respect to the Frobenius inner product $\langle\cdot,\cdot\rangle$. Thus we can define $\dot{Q}_i(t)$ as

$$\dot{Q}_i(t) = M_i(t) - \langle M_i(t), Q_i(t)\rangle Q_i(t) \tag{3.8}$$

for any $M_i(t) \in \mathbb{R}^{2\times 2}$.

In order to guarantee the increasing monotonicity of the real part of the rightmost eigenvalue, we construct a scaled gradient flow on $\mathcal{S}_F$, maximizing the derivative of the real part of the rightmost eigenvalue using the expression provided by Lemma 3.1.

Let us consider again for a simple rightmost eigenvalue λ and its left and right eigenvectors x, y the matrices X, Y and $\Gamma_i, i = 1, \ldots, m$ defined as in Theorem 3.3, that is

$$X = [\Re(x)\ \Im(x)],\ Y = [\Re(y)\ \Im(y)],\ \Gamma_i = \begin{bmatrix} \Re(p_i(\lambda)) & -\Im(p_i(\lambda)) \\ \Im(p_i(\lambda)) & \Re(p_i(\lambda)) \end{bmatrix}.$$

Neglecting for sake of conciseness the parameter t and letting

$$\xi = x^*\left(\sum_{i=1}^{m}\Big(A_i - \frac{\epsilon}{w_i}U_iQ_iV_i^T\Big)\right)y > 0,$$

we have

$$\begin{aligned}\Re(\dot\lambda) &= \textstyle\sum_{i=1}^m \frac{\epsilon}{\xi w_i}\Re\big(x^*(\dot U_i Q_i V_i^T + U_i\dot Q_i V_i^T + U_i Q_i \dot V_i^T)y\big)p_i(\lambda)\\ &= \textstyle\sum_{i=1}^m \frac{\epsilon}{\xi w_i}\big\langle X, (I_n - U_iU_i^T)R_iQ_iV_i^TY\Gamma_i^T\big\rangle + \big\langle X, U_i(M_i - \langle M_i, Q_i\rangle Q_i)V_i^TY\Gamma_i^T\big\rangle\\ &\quad +\big\langle X, U_iQ_iS_i^T(I_n - V_iV_i^T)Y\Gamma_i^T\big\rangle\\ &= \textstyle\sum_{i=1}^m \frac{\epsilon}{\xi w_i}\big\langle X\Gamma_iY^TV_iQ_i^T, (I_n - U_iU_i^T)R_i\big\rangle + \big\langle U_i^TX\Gamma_iY^TV_i, M_i - \langle M_i, Q_i\rangle Q_i\big\rangle\\ &\quad +\big\langle S_i, (I_n - V_iV_i^T)Y\Gamma_i^TX^TU_iQ_i\big\rangle.\end{aligned}$$

Since $(I_n - U_iU_i^T), (I_n - V_iV_i^T)$ are semi-positive definite, an ascent direction is obtained by the following choices

$$R_i = X\Gamma_iY^TV_iQ_i^T,\ M_i = U_i^TX\Gamma_iY^TV_i,\ S_i = Y\Gamma_i^TX^TU_iQ_i.$$

Inserting these choices in (3.7) and (3.8) brings us to the differential equations

$$\begin{cases}\dot U_i = \Big(I_n - U_iU_i^T\Big)X\Gamma_iY^TV_iQ_i^T,\\ \dot V_i = \Big(I_n - V_iV_i^T\Big)Y\Gamma_i^TX^TU_iQ_i,\\ \dot Q_i = U_i^TX\Gamma_iY^TV_i - \langle U_i^TX\Gamma_iY^TV_i, Q_i\rangle Q_i.\end{cases}\tag{3.9}$$

Finally, we can also exploit the fact that the critical perturbations share the same space of columns and rows. This result allows us to use a lower number of variables, thus less memory space. Therefore, we restrict to the sets of perturbations that share same column and row spaces. We move along a path defined by the compact singular value decomposition of the kind

$$\delta A_i(t) = -\frac{\epsilon}{w_i}U(t)Q_i(t)V(t),\ i = 1,\dots,m$$

where $U(t), V(t) \in \mathbb{R}^{n\times 2}$ are such that $U(t)^T\dot U(t) = V(t)^T\dot V(t) = 0\ \forall\ t \geq 0$ and $Q_i(t) \in \mathbb{R}^{2\times 2}$ s.t. $\|Q_i(t)\|_F = 1,\ \forall\ t \geq 0,\ i = 1,\dots,m$. Through steps analogous to the previous ones, we get to the derivatives

$$\begin{cases}\dot U = \Big(I_n - UU^T\Big)X\frac{\sum_{i=1}^m \frac{1}{w_i}\Gamma_iY^TVQ_i^T}{m},\\ \dot V = \Big(I_n - VV^T\Big)Y\frac{\sum_{i=1}^m \frac{1}{w_i}\Gamma_i^TX^TUQ_i}{m},\\ \dot Q_i = U^TX\Gamma_iY^TV - \langle U^TX\Gamma_iY^TV, Q_i\rangle Q_i.\end{cases}\tag{3.10}$$

Remark 3.5. *For the special case of the standard eigenvalue problem, the difference with [4] is as follows. For the case of a complex rightmost point a differential equation is proposed in [4] on the manifold of matrices of rank exactly equal to two, which is a Riemannian manifold, equipped with the natural geometry (see also [8]). In this way geodesics can be infinitely extended on the manifold, and the induced distance to a rank one matrix is infinite. Although geometrically very neat, the consequence is that distinction needs to be make in the overall algorithm between the case of a real valued and a complex valued rightmost point in the pseudospectra (rank one or rank two critical perturbations). In our setting we do not impose any rank constraints on Q_i, hence, we can address both cases simultaneously. In several examples we have even observed switches, in the sense that starting from a complex eigenvalue, the real rightmost point of the pseudospectrum is found.*

3.2 Implementation

In Algorithm 1 we present a basic description for the considered case of unstructured perturbations, based on discretizing (3.10) using the forward Euler method with adaptive stepsize, thereby ensuring monotonicity of the real part of the rightmost eigenvalue. Thus, we exploit the fact that perturbations $\delta A_i,\ i = 1, \ldots, m$ share the same space of columns and rows and we will move along a trajectory in the manifold $\mathcal{S}_F$, defined in (3.6). At each step, after the computation of the rightmost eigenvalue of the perturbed problem and its corresponding left and right eigenvectors, a normalization of the eigenvectors is needed in order to satisfy condition (3.3), and a normalization of the perturbations to ensure the next iterate lies in $\mathcal{S}_F$.

The initialization is as follows. We compute the rightmost eigenvalue of the unperturbed problem, and its corresponding left and right eigenvectors. Then we initialize matrices $X, Y, \Gamma_i,\ i = 1, \ldots, m$ as in Theorem 3.3. We apply a compact QR decomposition of matrices X, Y such that $X = Q_X R_X,\ Y = Q_Y R_Y$ and we define

$$U^{(1)} = Q_X,\ V^{(1)} = Q_Y,\ Q_i^{(1)} = \frac{R_X \Gamma_i R_Y^T}{\left\| R_X \Gamma_i R_Y^T \right\|_F},\ i = 1, \ldots, m.$$

Finally we construct the perturbed system $\sum_{i=1}^{m} \big(A_i - \frac{\epsilon}{w_i} U^{(1)} Q_i^{(1)} V^{(1)T}\big) p_i(\lambda)$ and, once again, we compute the rightmost eigenvalue and the corresponding eigenvectors. In order to avoid convergence to a locally rightmost point of the pseudospectrum which is not globally rightmost, Algorithm 1 can be repeated using different rightmost eigenvalues of the original problem as starting value.

4 Spectral norm

4.1 Foundations of the algorithm

The first theorem we state characterizes the optimal perturbations in terms of rank; since the proof is similar to the linear eigenvalue problem [4] the complete proof of this result can be found in appendix.

Theorem 4.1. *Let λ_{RM} be a globally rightmost point of the ϵ-pseudospectrum. Assume λ_{RM} is an eigenvalue for some ϵ-bounded perturbation $(\delta A_1, \ldots, \delta A_{m,})$, with*

$$\delta A_i = -\frac{\epsilon}{w_i} E_i, \quad \|E_i\|_2 \leq 1 \quad \textit{for all } i,$$

and x and y corresponding left and right eigenvectors normalized such that

$$x^* \left(\sum_{i=1}^{m} (A_i + \delta A_i) p_i'(\lambda_{\mathrm{RM}}) \right) y > 0. \tag{4.1}$$

Then $\|E_i\|_2 = 1,\ i = 1, \ldots, m$. Moreover, if λ_{RM} is a simple eigenvalue, then there exist matrices $\tilde{E}_i,\ i = 1, \ldots, m$, with unit spectral norm, such that

$$\left(\sum_{i=1}^{m} \left(A_i + \widetilde{\delta A_i} \right) p_i(\lambda_{\mathrm{RM}}) \right) y = 0, \quad x^* \left(\sum_{i=1}^{m} \left(A_i + \widetilde{\delta A_i} \right) p_i(\lambda_{\mathrm{RM}}) \right) = 0$$

where $\widetilde{\delta A_i} = -\frac{\epsilon}{w_i} \tilde{E}_i$ and $\tilde{E}_i$ has the following property:

(i) if λ_{RM} is real $\tilde{E}_i$ has rank-1;

Data: $U^{(1)}, Q_i^{(1)}, V^{(1)}, \lambda_1, x_1, y_1, r_1$
Result: $U^{(n+1)}, Q_i^{(n+1)}, V^{(n+1)}, \lambda_{n+1}, x_{n+1}, y_{n+1}, r_{n+1}$, $\alpha = \Re(\lambda_{n+1})$
begin

1. Set $n = 1$ and $\delta = r_1$
2. Compute
$$X = [\Re(x_n)\ \Im(x_n)],\ Y = [\Re(y_n)\ \Im(y_n)],\ \Gamma_i = \begin{bmatrix} \Re(p_i(\lambda_n)) & -\Im(p_i(\lambda_n)) \\ \Im(p_i(\lambda_n)) & \Re(p_i(\lambda_n)) \end{bmatrix},\ i = 1 \dots m$$
3. Compute the derivatives
$$\begin{cases} \dot{U} = \left(I_n - U^{(n)}U^{(n)T}\right) X \frac{\sum_{i=1}^m \frac{1}{w_i}\Gamma_i Y^T V^{(n)} Q_i^{(n)T}}{m} \\ \dot{V} = \left(I_n - V^{(n)}V^{(n)T}\right) Y \frac{\sum_{i=1}^m \frac{1}{w_i}\Gamma_i^T X^T U^{(n)} Q_i^{(n)}}{m} \\ \dot{Q}_i = U^{(n)T} X \Gamma_i Y^T V^{(n)} - \langle U^{(n)T} X \Gamma_i Y^T V^{(n)}, Q_i^{(n)} \rangle Q_i^{(n)} \end{cases}$$
4. Compute
$$\widehat{U}^{(n+1)} = U^{(n)} + r_n \delta \dot{U},\ \widehat{V}^{(n+1)} = V_{(n)} + r_n \delta \dot{V},\ \widehat{Q}_i^{(n+1)} = Q_i^{(n)} + r_n \delta \dot{Q}_i,\ i = 1, \dots, m.$$
5. Project these matrices on the original trajectory by imposing the properties discussed in the previous section; we denote with $\mathrm{Orth}(\widehat{U}^{(n+1)})$, $\mathrm{Orth}(\widehat{V}^{(n+1)})$ the matrices with orthonormal columns obtained by a compact QR decomposition of $\widehat{U}^{(n+1)}$, $\widehat{V}^{(n+1)}$ respectively:
$$U^{(n+1)} = \mathrm{Orth}(\widehat{U}^{(n+1)}),\ V^{(n+1)} = \mathrm{Orth}(\widehat{V}^{(n+1)}),\ Q_i^{(n+1)} = \frac{\widehat{Q}_i^{(n+1)}}{\left\|\widehat{Q}_i^{(n+1)}\right\|_F},\ i = 1, \dots, m.$$
6. Define $\delta A_i = -\frac{\epsilon}{w_i} U^{(n+1)} Q_i^{(n+1)} V^{(n+1)T}$, $i = 1, \dots, m$ and the new perturbed problem $\sum_{i=1}^m (A_i + \delta A_i) p_i(\lambda)$. Compute the rightmost eigenvalue $\hat{\lambda}$ of the new problem, and the associated left and right eigenvectors $\hat{x}, \hat{y}$.
7. **if** $\Re(\hat{\lambda}) \leq \Re(\lambda_n)$ **then**

   reject the step, compute a new step size $r := r_n/h$ and repeat starting from 4. We find a new $\bar{\lambda}$ and its corresponding left and right eigenvectors $\bar{x}, \bar{y}$.
8. **if** $r < r_n$ **then**

   we accept the step size r and set $\lambda_{n+1} = \bar{\lambda}, x_{n+1} = \bar{x}, y_{n+1} = \bar{y}, r_{n+1} = r$

   else

   compute another step size $\tilde{r} = h r_n$ and repeat starting from 4. We find a new $\tilde{\lambda}$ and its corresponding left and right eigenvectors $\tilde{x}, \tilde{y}$.

   if $\Re(\hat{\lambda}) \leq \Re(\tilde{\lambda})$ **then**

   we accept the step size $\tilde{r}$ and set $\lambda_{n+1} = \tilde{\lambda}, x_{n+1} = \tilde{x}, y_{n+1} = \tilde{y}, r_{n+1} = \tilde{r}$

   else

   we accept the step size r and set $\lambda_{n+1} = \hat{\lambda}, x_{n+1} = \hat{x}, y_{n+1} = \hat{y}, r_{n+1} = r_n$
9. Compute the Frobenius norm of the gradient of $\Re(\lambda_{n+1})$ with respect to the elements of matrices δA_i. If the norm is larger than the given tolerance set $n \leftarrow n+1$ and go to step 2., otherwise stop the algorithm.

Algorithm 1: Computation of the real pseudospectral abscissa, Frobenius norm case. Number $h > 1$ is the scaling factor for the adaptation of the stepsize r_n.

(ii) if λ_{RM} is not real $\tilde{E}_i$ has rank-2 and two singular values equal to 1.

Sketch of the proof. Let us consider the case when λ_{RM} is genuinely complex. First we prove that the optimal perturbations $\frac{\epsilon}{w_i}E_i$ are such that matrices E_i have spectral norm equal to 1. Subsequently, we show that the critical matrices share the column and row space respectively with X and Y^T, by considering a path through the critical perturbations and applying local optimality conditions. Considering a path $UR_i(t)V^T$ with $U, V \in \mathbb{R}^{n\times 2}$ spanning the column space of X and Y respectively, and $R_i(t) \in \mathbb{R}^{2\times 2}$, we finally prove that $R_i(0)$ is orthogonal, so that the optimal E_i can be expressed as a rank-two matrix with spectral norm equal to 1. □

Next we provide a characterization of rank-2 extremizers. For the proof we refer, once again, to the appendix.

Theorem 4.2. *Under the assumptions of Theorem 4.1, let λ_{RM} be a globally rightmost point of the ϵ-pseudospectrum with λ_{RM} a simple non-real eigenvalue. Let x, y be its left and right eigenvectors, and let matrices X, Y, Γ_i, $i = 1, \ldots, m$ be defined as in Theorem 3.3.*

If $\delta A_i = -\frac{\epsilon}{w_i}E_i$ has rank-2 for $i = 1, \ldots, m$, then

$$E_i = UQ_iV^T \tag{4.2}$$

where $U, V \in \mathbb{R}^{n\times 2}$ have orthonormal columns and $Q_i \in \mathbb{R}^{2\times 2}$ with

(i) Q_i is an orthogonal matrix;

(ii) X and U have the same range, Y and V have the same range;

(iii) the matrix $Q_i^T U^T X \Gamma_i^T Y^T V$ is symmetric and positive definite, for $i = 1, \ldots, m$.

The following theorem is a main theoretical result regarding the spectral norm case; it provides an explicit characterization of the optimal perturbation in case of either a real or a genuine complex rightmost eigenvalue of the pseudospectrum.

Theorem 4.3. *Let λ_{RM} be a globally rightmost point of ϵ-pseudospectruum.*
Case 1: λ_{RM} is non-real. *Assume λ_{RM} is a simple eigenvalue for some ϵ-size perturbation $(\delta A_1, \ldots, \delta A_m)$, where each δA_i has rank two and two singular values equal to one, and let x and y be the corresponding left and right eigenvectors normalized such that*

$$x^* \left(\sum_{i=1}^{m} (A_i + \delta A_i) p_i'(\lambda_{\mathrm{RM}}) \right) y > 0.$$

Let X, Y and Γ_i be defined as in Theorem 3.3. Let the decomposition

$$X\Gamma_i Y^T = U_i \Sigma_i V_i^T, \quad i = 1, \ldots, m, \tag{4.3}$$

correspond to the compact singular value decomposition. Then we have

$$\delta A_i = -\frac{\epsilon}{w_i} U_i V_i^T, \ i = 1, \ldots, m.$$

Case 2: λ_{RM} is real. *Assume λ_{RM} is a simple eigenvalue for some ϵ-size perturbation $(\delta A_1, \ldots, \delta A_m)$, where each δA_i has rank one, and let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ be corresponding left and right eigenvectors normalized such that*

$$x^* \left(\sum_{i=1}^{m} (A_i + \delta A_i) p_i'(\lambda_{\mathrm{RM}}) \right) y > 0.$$

then we have

$$\delta A_i = -\frac{\epsilon}{w_i}\text{sign}(p_i(\lambda_{\text{RM}}))\ xy^T,\ i = 1, \ldots, m. \tag{4.4}$$

Proof. We restrict to the first case. For the case of a rightmost real eigenvalue the derivation is the same as for complex pseudospectra and we refer to [14]. By a property of the singular value decomposition and the fact that Γ_i is invertible, X and U_i have the same column space, and Y and V_i have the same column space. Based on the first two properties in Theorem 4.2 we can write

$$\delta A_i = -\frac{\epsilon}{w_i}U_i Q_i V_i^T, \tag{4.5}$$

where Q_i is some 2-by-2 orthogonal matrix for $i = 1, \ldots, m$.

From the last property in Theorem 4.2 we have

$$Q_i^T U_i^T X \Gamma_i Y^T V_i > 0,\ i = 1, \ldots, m,$$

which, using (4.3), can be rewritten as

$$Q_i^T \Sigma_i > 0,\ i = 1, \ldots, m. \tag{4.6}$$

Assuming for the moment that Q_i corresponds to a rotation, i.e.

$$Q_i = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix},\ \theta \in (-\pi,\ \pi],$$

and letting $\Sigma = \text{diag}(\sigma_1, \sigma_2)$, then the satisfaction of condition (4.6) requires $\theta = 0$. Using a similar argument one can show that Q_i cannot correspond to a reflection. In conclusion, the only possibility is to have $Q_i = I,\ i = 1, \ldots, m$, in (4.5). □

Besides the provision of explicit optimality conditions, the importance of Theorem 4.3 is that it directly gives rise to a fixed-point iteration for computing the pseudospectral abscissa, thereby generalizing the method of [14] from complex to real perturbations, and the method of [17] from linear to nonlinear eigenvalue problems. However, as we have done with the Frobenius norm case, we construct a scaled gradient flow for a functional defined in the perturbations space, which has the advantage of guaranteeing monotonicity. Consider a path in the space

$$\begin{aligned}\mathcal{S}_S := \big\{(\delta A_1, \ldots, \delta A_m) \in \mathbb{R}^{n\times n\times m} : \text{rank}(\delta A_i) = 2 \text{ with nonzero singular values} \\ \text{equal to ech other, and } w_i \left\|\delta A_i\right\|_2 = \epsilon,\ i = 1, \ldots, m\big\},\end{aligned} \tag{4.7}$$

parametrized as

$$\delta A_i(t) = -\frac{\epsilon}{w_i}U_i(t) Q_i(t) V_i(t)^T$$

with $U_i(t),\ V_i(t) \in \mathbb{R}^{n\times 2}$ s.t. $U_i(t)^T \dot{U}_i(t) = V_i(t)^T \dot{V}_i(t) = 0$ and $Q_i(t) \in \mathbb{R}^{2\times 2}$ orthogonal. The condition on $Q_i(t)$ implies that matrix $Q_i(t)^T \dot{Q}_i(t)$ is skew-symmetric. In order to guarantee monotonicity to our iterations, we impose again the derivative of the real part of the rightmost eigenvalue to be nonnegative. With analogous steps to the Frobenius norm case, we get the same expressions in (3.9) for the derivatives $\dot{U}_i, \dot{V}_i$, whereas the new constraint of orthogonality on Q_i leads to the following derivative

$$\dot{Q}_i = Q_i\ \text{skew}(Q_i^T U_i^T X \Gamma_i Y^T V_i), \tag{4.8}$$

with skew$(P) = (P - P^T)/2$. For more details we refer to [4], where the methodology is fully worked for the standard eigenvalue problem. However, in our implementation we exploit, in addition, the fact that optimal perturbations δA_i share the same column and row spaces in order to save memory space; in this case we will use expressions in (3.10) for derivatives $\dot{U}, \dot{V}$ and the same formula in (4.8) for $\dot{Q}_i$.

Remark 4.4. *As for the Frobenius norm case, this construction allows us to treat simultaneously the cases of either a real or complex globally rightmost point of the pseudospectrum. We know from [14] that in case of a real rightmost point, the optimal perturbations can be expressed as rank-one matrices; however, thanks to the properties of the spectral norm, the same eigenvalue can be obtained through perturbation matrices with a higher rank and the same largest singular value. For this reason our algorithm is able to converge to a real rightmost point moving along a path of rank-two matrices, where the orthogonality of matrices Q_i is imposed. Indeed, in several simulations we could observe our algorithm converging to a rightmost real eigenvalue although it started from a genuinely complex value.*

4.2 Implementation

The algorithm is developed analogously to Algorithm 1, where the only main differences are given by the new differential equation for matrices Q_i and the re-orthogonalization after every step. The characterization of the optimal perturbations given in Theorem 4.3 is exploited in the initialization of the algorithm, as we clarify in what follows.

Let us first assume that the starting eigenvalue is complex. If XY^T has the compact singular value decomposition $XY^T = U\Sigma V^T$ and if

$$(U^T X)\Gamma_i(Y^T V) = \tilde{U}_i \Sigma_i \tilde{V}_i^T \in \mathbb{R}^{2\times 2}$$

corresponds, once again, to a singular value decomposition, then the decomposition (4.3) holds, where

$$U_i = U\tilde{U}_i, \quad V_i = V\tilde{V}_i, \ i = 1, \dots, m.$$

This has two consequences. First, to compute the singular value decomposition of $X\Gamma_i Y^T$, $i = 1, \dots, m$ only one compact singular value decomposition of a n-by-n matrix needs to be performed. Second, the optimal perturbations have the form

$$\delta A_i = -\frac{\epsilon}{w_i} U \left(\tilde{U}_i \tilde{V}_i^T\right) V^T, \ i = 1, \dots, m.$$

From the starting eigenvalue and corresponding eigenvectors we can compute X, Y and Γ_i, $i = 1, \dots, m$. Subsequently, the gradient flow discretization is initialized with U, V and $Q_i(0) := \tilde{U}_i \tilde{V}_i^T$. Note that his initialization can be interpreted as performing one iteration of the fixed-point iteration as in [17].

If the starting eigenvalue λ is real, we obtain U and V by extending eigenvectors x and y to $n \times 2$ matrices with orthogonal columns, and initialize Q_i with $\text{sign}(p_i(\lambda)) \otimes I_2$.

Remark 4.5. *In case Assumption 2.2 is not satisfied for the rightmost point of the pseudospectrum (i.e., a complex rightmost point with real eigenvector), the critical perturbations have rank one for the Frobenius norm case and can be chosen as rank one for the spectral norm case. In such a case our algorithms can still converge to the correct value, because of the arguments spelled out in Remark 3.5 and Remark 4.4.*

5 Numerical experiments

In this section we show the results obtained from a MATLAB implementation of the algorithms previously described, illustrating applications to the delay eigenvalue problem and to the polynomial eigenvalue problem, and using the Frobenius and spectral norm. For the delay eigenvalue problem, rightmost eigenvalues are computed with the method of [12] for systems of delay differential algebraic equations, generalizing the approach of [22]. We start with an introductory example.

Example 1. Consider the delay eigenvalue problem

$$\Big(\lambda I - (A_1 + \delta A_1) - (A_2 + \delta A_2)e^{-\lambda\tau}\Big)y = 0 \tag{5.1}$$

with

$$A_1 = \begin{bmatrix} -5 & 1 \\ 2 & -6 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} -2 & 1 \\ 4 & -1 \end{bmatrix}, \qquad \tau = 1$$

and the weights vector $w = [1, +\infty]$; thus, we allow a perturbation only on matrix A_1. In Figure 5.1 the complex and real pseudospectra are shown for the spectral norm and for $\epsilon = 3.5$. The pseudospectra contours are computed from a singular value characterization using a grid in the complex plane. For the complex pseudospectrum the singular value characterization can be found in [13]. For real pseudospectra we have, since only one matrix is perturbed, the expression

$$\Lambda_\epsilon = \left\{\lambda \in \mathbb{C} : \ \mu_{\mathbb{R}}\left((\lambda I - A_0 - A_1 e^{-\lambda\tau})^{-1}\right) \geq \epsilon^{-1}\right\},$$

with $\mu_{\mathbb{R}}(\cdot)$ the real structured singular value, for which a computational expression is derived in [16].

In this particular case the pseudospectral abscissa is the same, since it is located on the real axis, where the two pseudospectra always coincide. However, a clear difference between the two sets is observable, which illustrates the powerful utility of our work. The importance of using several rightmost eigenvalues as starting values is also clarified from this figure; in this case, the iterations starting from the rightmost eigenvalues of the original unperturbed eigenvalue problem converge to a locally rightmost point of the pseudospectra, whereas the real pseudospectral abscissa is reached by starting from different original eigenvalues.

5.1 Experiments on the Frobenius norm

We show here the results obtained applying our algorithm to some of the delay eigenvalue problems in the benchmark collection used in [14], that appear as

$$\left(\lambda I_n - (B_0 + \delta B_0) - \sum_{i=1}^{l}(B_i + \delta B_i)e^{-\lambda\tau_i}\right)y = 0. \tag{5.2}$$

The weights are chosen as

$$w_i = \|B_i\|_F^{-1}, \ \forall\ i = 0, \ldots, m,$$

in order to allow a maximal relative perturbation equal to ϵ on each matrix (note that the first term is not perturbed, implying that the corresponding weight is ∞). In Table 5.1 we display the pseudospectral abscissa found in a few problems using different values for ϵ. The first column refers to the number of the problem in the collection.[1]

[1]The collection is available at the webpage http://twr.cs.kuleuven.be/research/software/delay-control/benchmark.

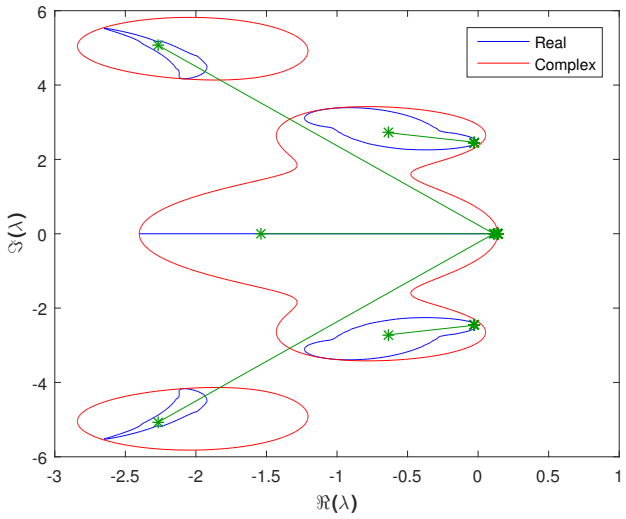

Figure 5.1: Comparison between the real and the complex pseudospectra of problem (5.1). The initial and final eigenvalue for iterations executed by the algorithm starting from different rightmost eigenvalues are represented by the green lines.

Table 5.1: Pseudospectral abscissa results obtained using the Frobenius norm as a perturbation measure. In the second column n refers to the problem dimension. The last column shows the number of iterations needed to obtain a ten-digit accuracy.

Prob.	$(n$, #delays)	α	ϵ	α_ϵ	# it (10 digits)
1	$(3, 1)$	$-2.866038425\text{e-}02$	1.e-1	$7.242821759\text{e-}02$	14
			1.e-2	$-1.463286850\text{e-}02$	9
			1.e-4	$-2.851134502\text{e-}02$	3
6	$(10, 7)$	$-3.775473572\text{e-}01$	1.e-1	$1.972654297+00$	329
			1.e-2	$-1.823559005\text{e-}01$	20
			1.e-4	$-3.754178862\text{e-}01$	4
10	$(4, 3)$	$-9.858488139\text{e-}02$	1.e-1	$-3.048429971\text{e-}02$	37
			1.e-2	$-9.189037810\text{e-}02$	19
			1.e-4	$-9.851861147\text{e-}02$	2

Table 5.2: Pseudospectral abscissa of the polynomial eigenvalue problem obtained using the Frobenius norm as a perturbation measure.

ϵ	α	α_ϵ	# it (10 digits)
10^{-1}	9.462649021e-02	1.649534804e-01	49
$10^{-1.5}$		1.160533627e-01	30
10^{-2}		1.013171374e-01	10
$10^{-2.5}$		9.673361108e-02	4
10^{-3}		9.529195135e-02	2
10^{-4}		9.469300010e-02	2

Table 5.3: The same problems of table 5.1 are investigated here, using the spectral norm as a perturbation measure.

Prob.	$(n$, #delays)	α	ϵ	α_ϵ	# it (10 digits)
1	$(3, 1)$	-2.866038425e-02	1.e-1	5.607794071e-02	22
			1.e-2	-1.716339042e-02	5
			1.e-4	-2.853917009e-02	2
6	$(10, 7)$	-3.775473572e-01	1.e-1	1.019638143+00	305
			1.e-2	-2.307810008e-01	19
			1.e-4	-3.759759125e-01	3
10	$(4, 3)$	-9.858488139e-02	1.e-1	-2.026579142e-02	32
			1.e-2	-9.402496484e-02	15
			1.e-4	-9.853953912e-02	2

As already discussed, the algorithm can also be applied to polynomial eigenvalue problems with the form

$$\left(\sum_{i=1}^{m}(A_i + \delta A_i)\lambda^{i-1}\right)y = 0. \tag{5.3}$$

Here we consider the same quadratic eigenvalue problem used in [14], where

$$A_1 = \begin{bmatrix} 121 & 18.9 & 15.9 \\ 0 & 2.7 & 0.145 \\ 11.9 & 3.64 & 15.5 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 7.66 & 2.45 & 2.1 \\ 0.23 & 1.04 & 0.223 \\ 0.6 & 0.756 & 0.658 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 17.6 & 1.28 & 2.89 \\ 1.28 & 0.824 & 0.412 \\ 2.89 & 0.413 & 0.725 \end{bmatrix}.$$

The pseudospectral abscissas obtained for different values of ϵ are displayed in Table 5.2; in each experiment unitary weights were applied.

5.2 Experiments on the spectral norm

We give in this subsection the results produced by the algorithm on the same delay eigenvalue problems previously introduced using the spectral norm instead of the Frobenius one. Analogously to the previous case, we set the weights

$$w_i = \|B_i\|_2^{-1}, \ \forall\ i = 0, \ldots, m.$$

The results are resumed in Table 5.3. Finally, the same quadratic eigenvalue problem introduced in the previous subsection is investigated by means of the spectral norm as perturbation measure; the pseudospectral abscissas obtained thanks to our algorithm for different ϵ and unitary weights are reported in Table 5.4.

Table 5.4: Pseudospectral abscissa of the polynomial eigenvalue problem obtained using the spectral norm as a perturbation measure.

ϵ	α	α_ϵ	# it (10 digits)
10^{-1}	9.462649021e-02	1.749188888e-01	25
$10^{-1.5}$		1.188936039e-01	15
10^{-2}		1.021872550e-01	9
$10^{-2.5}$		9.700605828e-02	3
10^{-3}		9.537783897e-02	2
10^{-4}		9.470157776e-02	2

The choice of the weights for these simulations was made again with the aim of allowing a maximal relative perturbation of magnitude ϵ. Hence, the weights used for the Frobenius norm case and spectral norm case were different; however, if we had used the same weights, we would have noticed that the pseudospectral abscissa produced by the spectral norm is larger or equal to the one obtained using the Frobenius norm. This is because the set of the unitary Frobenius norm matrices is included in the set of the unitary spectral norm matrices. Moreover, using the same weights, the values found with the two different norms would have coincided whether the rightmost point of the pseudospectrum had been real; this is easily explained by the fact that the optimal perturbations can be expressed as a set of rank-one matrices, and in this case the spectral and the Frobenius norm coincide.

6 Extensions of the Frobenius norm case

6.1 Structured real perturbations

In this section we will briefly discuss how to incorporate a particular structure in the set of real perturbations of the original non linear eigenvalue problem. Namely we study the perturbed problem

$$\left(\sum_{i=1}^{m}(A_i + D_i\delta A_i(t)E_i)p_i(\lambda)\right)y = 0, \tag{6.1}$$

where D_i and E_i are real valued shape matrices of appropriate dimensions. The dimensions of the perturbations δA_i may differ from one coefficient matrix to another. The perturbation measure defining the corresponding real structured ϵ-pseudospectrum is given by (2.1).

Approach

The following theorem, analogously to Theorem 3.3, provides explicit optimality conditions, which characterize the critical set of rank-two perturbations corresponding to a globally rightmost point.

Theorem 6.1. *Let λ_{RM} be a globally rightmost point in the structured ϵ-pseudospectrum and assume it is a simple eigenvalue for perturbations $(\delta A_1, \ldots, \delta A_m)$. Let x, y be its corresponding left and right eigenvectors normalized such that*

$$x^*\left(\sum_{i=1}^{m}(A_i + D_i\delta A_i E_i)p_i'(\lambda_{\mathrm{RM}})\right)y > 0.$$

Let X, Y and Γ_i, $i = 1, \ldots, m$ be defined as in Theorem 3.3. Then for each $i = 1, \ldots, m$, $D_i^T X \Gamma_i Y^T E_i^T$ can be either null or non-null. In the latter case, the optimal perturbations can be expressed as

$$\delta A_i = -\frac{\epsilon}{w_i} \frac{D_i^T X \Gamma_i Y^T E_i^T}{\left\| D_i^T X \Gamma_i Y^T E_i^T \right\|_F}.$$

Proof. The proof is similar to the proof of Theorem 3.3. Let us consider again the ball of radius ϵ containing the perturbation sets $(\delta A_1, \ldots, \delta A_m)$ such that $\|\Delta\|_{\text{glob}} \leq \epsilon$. As in Theorem 3.3, for each $i \in \{1, \ldots, m\}$, if the optimum set lies on the border of the ϵ-ball, the gradient $\frac{\partial \Re(\lambda_{\text{RM}})}{\partial \delta A_i}$ has the same direction and is oriented as δA_i itself. However, unlike to the unstructured case we cannot exclude that this gradient is zero. □

As we shall illustrate, with structured perturbations it is possible that some of the critical perturbations δA_i are internal points of the (weighted) ϵ-ball around A_i; therefore, instead of looking for an optimum set on the boundary, we will construct a trajectory that possibly passes through the internal part of the ball. The following proposition shows that in case of a null gradient in an internal point, the optimum can still be reached when restricting to up to rank-two perturbations.

Proposition 6.2. *Let λ_{RM} be a globally rightmost point of the structured ϵ-pseudospectrum and assume it is an eigenvalue for ϵ-bounded perturbation $(\delta A_1, \ldots, \delta A_m)$. Then there exists an ϵ-bounded perturbation $(\delta A_1^*, \ldots, \delta A_m^*)$, where each matrix δA_i^* has rank at most two, such that the eigenvalue λ_{RM} is preserved.*

Proof. Suppose, without loss of generality that $E_i \in \mathbb{R}^{q_i \times n}$. For every $i = 1, \ldots, m$ we define a matrix H_i such that the first two columns' span includes the space spanned by $\Re(E_i y)$ and $\Im(E_i y)$, and the other columns complete the basis of $\mathbb{R}^{q_i \times q_i}$, so that H_i is orthogonal. Thus, we can write $\delta A_i = \widetilde{\delta A_i} H_i^T$, $i = 1, \ldots, m$. With this factorization,

$$0 = \left(\sum_{i=1}^{m} (A_i + D_i \delta A_i E_i) p_i(\lambda_{\text{RM}}) \right) y = \sum_{i=1}^{m} (A_i + D_i \widetilde{\delta A_i} H_i^T) p_i(\lambda_{\text{RM}}) \Big(\Re(E_i y) + i \Im(E_i y) \Big) =$$

$$= \sum_{i=1}^{m} (A_i + D_i \widetilde{\delta A_i} p_i(\lambda_{\text{RM}})) \begin{pmatrix} \langle H_i^{(1)}, \Re(E_i y) + i \Im(E_i y) \rangle \\ \langle H_i^{(2)}, \Re(E_i y) + i \Im(E_i y) \rangle \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

From the last expression it is easy to see that we can get rid of the last $q - 2$ columns of H_i. Thus, for every $i = 1, \ldots, m$ we reset all elements in H_i to zero except for the elements in the first two columns. Therefore, defining $\delta A_i^* = \widetilde{\delta A_i} H_i$ we obtain at most rank-two matrices that preserve the λ_{RM} eigenvalue and its right eigenvector y. □

In the structured case we then need to redefine the set to which the trajectory in the space of perturbations is restricted, that is

$$\hat{\mathcal{S}}_F := \{(\delta A_1, \ldots, \delta A_m) : \operatorname{rank}(\delta A_i) \leq 2 \text{ and } w_i \|\delta A_i\|_F \leq \epsilon,\ i = 1, \ldots, m\}. \tag{6.2}$$

We use again a compact singular value decomposition as $\delta A_i(t) = -\frac{\epsilon}{w_i} U_i(t) Q_i(t) V_i(t)^T$, with $U_i(t), V_i(t) \in \mathbb{R}^{n \times 2}$ s.t. $U_i(t)^T \dot{U}_i(t) = V_i(t)^T \dot{V}_i(t) = 0$ and $Q_i(t) \in \mathbb{R}^{2 \times 2}$ s.t.$\|Q_i(t)\|_F \leq$

1, $i = 1, \dots, m$. Through the same steps seen in the unstructured case we state the following expressions for the derivatives

$$\begin{cases} \dot{U}_i = \Big(I_n - U_i U_i^T\Big) D_i^T X \Gamma_i Y^T E_i^T V_i Q_i^T, \\ \dot{V}_i = \Big(I_n - V_i V_i^T\Big) E_i Y \Gamma_i^T X^T D_i U_i Q_i, \\ \dot{Q}_i = \begin{cases} M_i - \langle M_i, Q_i \rangle Q_i, & \text{if } \|Q_i\|_F = 1,\ \langle M_i, Q_i \rangle > 0, \\ M_i, & \text{otherwise,} \end{cases} \end{cases} \tag{6.3}$$

with $M_i = U_i^T D_i^T X \Gamma_i Y^T E_i^T V_i$. In this way we ensure $\|Q_i\|_F \leq 1$. We only project gradient M_i if $\|Q_i\|_F = 1$ and the gradient points outwards. The adaptations to Algorithm 1 are straightforward.

Numerical example

Let us consider the following perturbed delay system

$$\dot{x}(t) = \Big(A + D\delta A E\Big) x(t - \tau) \tag{6.4}$$

where δA is a scalar, the delay τ is set equal to 0.1 and

$$A = \frac{1}{100} \begin{bmatrix} 0 & 4 & 0 & 0 \\ -301 & -56 & 301 & 224 \\ 0 & 0 & 0 & 16 \\ 1.9531 & 109.375 & -3.9063 & -437.5 \end{bmatrix}, \quad D = \begin{bmatrix} 0 \\ 3.125 \\ 0 \\ 0 \end{bmatrix}, \quad E = \begin{bmatrix} 1.6 & 0 & -1.6 & 0 \end{bmatrix}.$$

In figure 6.1 we represent the behaviour of the rightmost eigenvalue of system (6.4) as the perturbation scalar parameter δA varies between 0 and 0.4. It is easy to observe that if $\epsilon = 0.1$ (whose correspondent pseudospectral abscissa is 1.22788e−02) then the rightmost point of the pseudospectrum is obtained through a maximum norm perturbation; however, if for instance $\epsilon = 0.3$ (whose correspondent pseudospectral abscissa is 1.22939e−02), the pseudospectral abscissa is not generated anymore by a maximum norm perturbation. This example clarifies the reason why in the structured perturbation case we can't impose the norm of the matrices Q_i to be unitary. Given such an ϵ, our algorithm provides the rightmost point of the pseudospectrum and the associated critical perturbation, which in this case doesn't have a maximum norm.

6.2 New measure on the combined perturbations

Let us analyze a third case with a new perturbation measure instead of the one previously used. We use the Frobenius norm on the set constituted by all the m perturbation matrices; with this new definition, introducing again weights w_i, a set of perturbations is ϵ-bounded if $\|[w_1 \delta A_1 \ \cdots \ w_m \delta A_m]\|_F \leq \epsilon$, that is,

$$\sqrt{\sum_{i=1}^{m} w_i^2 \|\delta A_i\|_F^2} \leq \epsilon. \tag{6.5}$$

We are interested in computing the pseudospectral abscissa of the unstructured pseudospectrum; with this definition, having an optimum on the border of the weighted ϵ-ball around $(\delta A_1, \dots, \delta A_m)$ doesn't imply that the optimal individual perturbations δA_i have maximum norm, as in the two previous cases. Therefore, from the optimal perturbations set, we can assess the sensitivity, i.e., we can state which term mostly contributes to pushing the spectral abscissa to the right.

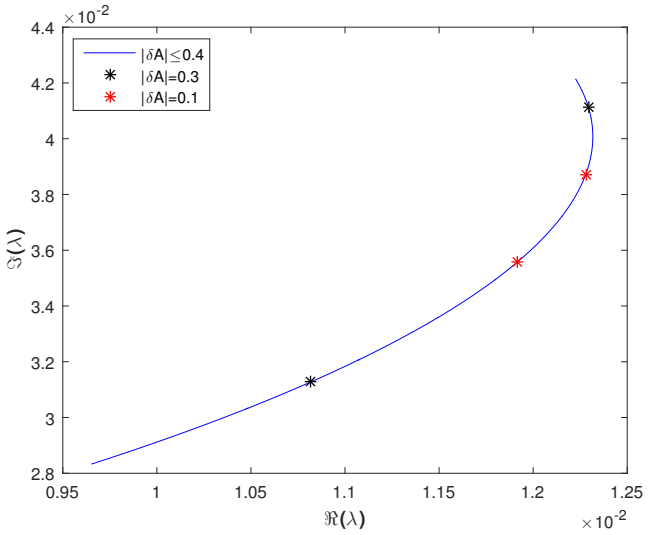

Figure 6.1: The root locus represents the rightmost eigenvalue of system (6.4) under different values for the scalar perturbation δA.

Approach

We start, once again, with the optimality conditions. Since the proof is analogous to the proof of Theorem 3.3, we omit it for sake of conciseness of the presentation.

Proposition 6.3. *Let λ_{RM} be a globally rightmost point in the ϵ-pseudospectrum corresponding to (6.5), and assume it is simple. Let x and y be its corresponding left and right eigenvectors normalized such that condition (3.3) holds. Defining X, Y, Γ_i, $i = 1, \dots, m$ as in Theorem 3.3, we can express the optimal perturbations as*

$$\delta A_i = -\frac{\epsilon}{w_i} \frac{X \Gamma_i Y^T}{\|(X\Gamma_1 Y^T, \dots, X\Gamma_m Y^T)\|_F}. \tag{6.6}$$

Also in this case our algorithm is based on a scaled gradient flow in a set of up to rank-two matrices, using a compact singular value decomposition $\delta A_i = -\frac{\epsilon}{w_i} U(t) Q_i(t) V(t)^T$, $i = 1, \dots, m$ where $U(t), V(t)$ have orthonormal columns and $\sqrt{\sum_{i=1}^m \|Q_i\|_F^2} = 1$. If we define the vector $Q = (Q_1, \dots, Q_m)$, this condition becomes $\|Q\|_F = 1$ and, following the same steps used before, we obtain the expressions for the derivatives

$$\begin{cases} \dot{U} = \left(I_n - UU^T\right) X \frac{\sum_{i=1}^m \frac{1}{w_i} \Gamma_i Y^T V Q_i^T}{m} \\ \dot{V} = \left(I_n - VV^T\right) Y \frac{\sum_{i=1}^m \frac{1}{w_i} \Gamma_i^T X^T U Q_i}{m} \\ \dot{Q} = M - \langle M, Q \rangle Q, \end{cases}$$

where $M = \left(\frac{1}{w_1} U^T X \Gamma_1 Y^T V, \dots, \frac{1}{w_m} U^T X \Gamma_m Y^T V\right)$.

Numerical example

We present the results obtained using the combined Frobenius norm as a perturbation measure for the same delay eigenvalue problems already investigated in Table 5.1. As previously done, we set weights $w_i = \|B_i\|_F^{-1}$, $i = 1, \dots, m$.

Table 6.1: Pseudospectral abscissa results obtained using the combined Frobenius norm as a perturbation measure. In the second column n refers to the problem dimension. The last two columns show the minimum and the maximum Frobenius norm assumed by the matrices Q_i in the optimum.

Prob.	(n, #delays)	α	ϵ	α_ϵ	$\min \|Q_i\|_F$	$\max \|Q_i\|_F$
1	$(3, 1)$	$-2.866038425\text{e-}02$	$\sqrt{2}$ 1.e-1	7.693994983e-02	0.434	0.901
			$\sqrt{2}$ 1.e-2	$-1.430507964\text{e-}02$	0.535	0.845
			$\sqrt{2}$ 1.e-4	$-2.850828236\text{e-}02$	0.552	0.834
6	$(10, 7)$	$-3.775473572\text{e-}01$	$\sqrt{8}$ 1.e-1	2.612751292+00	0.036	0.948
			$\sqrt{8}$ 1.e-2	$-1.532662982\text{e-}01$	0.223	0.778
			$\sqrt{8}$ 1.e-4	$-3.751418130\text{e-}01$	0.224	0.745
10	$(4, 3)$	$-9.858488139\text{e-}02$	$\sqrt{4}$ 1.e-1	1.231685411e-01	0.065	0.964
			$\sqrt{4}$ 1.e-2	$-9.072068875\text{e-}02$	0.176	0.798
			$\sqrt{4}$ 1.e-4	$-9.850744365\text{e-}02$	0.171	0.775

In order to compare the present case with the unstructured one equipped with the previous global norm, we set a maximum perturbation measure that include the case when all the matrices δA_i have norm equal to $\dfrac{\epsilon}{w_i}$. To this purpose, we adopt in each case the new maximum perturbation measure $\tilde{\epsilon} = \sqrt{m}\ \epsilon$, where m is the number of matrices δA_i. In Table 6.2 the results for these simulations are reported; as expected, we always find a spectral abscissa which is larger than the one obtained in the cases considered in section 5.1.

7 Conclusions

We presented various algorithms for the computation of real pseudospectral abscissa for a broad class of nonlinear eigenvalue problems. These iterative algorithms fully exploit the structure of the characteristic matrix and, possibly, additional structure on the coefficient matrices.

The approach is particularly appealing for nonlinear eigenvalue problems, since to the best of our knowledge no alternative approaches are available, unless one resorts to special cases (e.g., linear eigenvalue problem) or enlarges the class of perturbations. In addition, since low rank dynamics and sharing of column and row spaces significantly reduces the number of variables and memory cost in the gradient based scheme, in combination with iterative methods for selected eigenvalue computations (see, e.g., [7, 21] for generic nonlinear eigenvalue problems and [6] for the delay eigenvalue problem), the approach has potential for large scale problems.

Acknowledgments

The work of FB and WM was supported by the Programme of Interuniversity Attraction Poles of the Belgian Federal Science Policy Office (IAP P6-DYSCO), by OPTEC, the Optimization in Engineering Center of the KU Leuven, by the project G.0712.11N of the Research Foundation-Flanders (FWO - Vlaanderen), and by the project UCoCoS, funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 675080.

References

[1] R. Byers. A bisection method for measuring the distance of a stable matrix to the unstable matrices. *SIAM Journal on Scientific and Statistical Computing*, 9(9):875–881, 1988.

[2] M. Freitag and A. Spence. A Newton-based method for the calculation of the distance to instability. *Linear Algebra and its Applications*, 435(12):3189–3205, 2011.

[3] N. Guglielmi and C. Lubich. Differential equations for roaming pseudospectra: paths to extremal points and boundary tracking. *SIAM Journal on Numerical Analysis*, 49:1194–1209, 2012.

[4] N. Guglielmi and C. Lubich. Low-rank dynamics for computing extremal points of real pseudospectra. *SIAM Journal of Matrix Analysis and Applications*, 34:40–66, 2013.

[5] N. Guglielmi and M.L. Overton. Fast algorithms for the approximation of the pseudospectral abscissa and pseudospectral radius of a matrix. *SIAM Journal on Matrix Analysis and Applications*, 32:1166–1192, 2011.

[6] E. Jarlebring, K. Meerbergen, and W. Michiels. A Krylov method for the delay eigenvalue problem. *SIAM Journal on Scientific Computing*, 32(6):3278–3300, 2010.

[7] E. Jarlebring, W. Michiels, and K. Meerbergen. A linear eigenvalue algorithm for the nonlinear eigenvalue problem. *Numerische Mathematik*, 122(1), 2012.

[8] O. Koch and C. Lubich. Dynamical low rank approximation. *SIAM J. Matrix Anal. Appl.*, 29(2):434–454, 2007.

[9] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations: analysis and numerical solution*. Textbook in Mathematics. EMS Publishing House, 2006.

[10] K. Meerbergen, E. Mengi, W. Michiels, and R. Van Beeumen. Computation of pseudospectral abscissa for large scale nonlinear eigenvalue problems . *IMA Journal of Numerical Analysis*, 2016. Submitted for publication.

[11] E. Mengi, E.A. Yildirim, and M. Kilic. Numerical optimization of eigenvalues of Hermitian matrix functions. *SIAM Journal of Matrix Analysis and Applications*, 35(2):699–74, 2014.

[12] W. Michiels. Spectrum based stability analysis and stabilization of systems described by delay differential algebraic equations. *IET Control Theory and Applications*, 5(16):1829–1842, 2011.

[13] W. Michiels, K. Green, T. Wagenknecht, and S.-I. Niculescu. Pseudospectra and stability radii for analytic matrix functions with applications to time-delay systems. *Linear Algebra and its Applications*, 418(1):315–335, 2006.

[14] W. Michiels and N. Guglielmi. An iterative method for computing the pseudospectral abscissa for a class of nonlinear eigenvalue problems. *SIAM Journal on Scientific Computing*, 34(4):2366–2393, 2012.

[15] W. Michiels and S.-I. Niculescu. *Stability and stabilization of time-delay systems. An eigenvalue based approach*. SIAM, 2007.

[16] L. Qiu, B. Bernhardsson, A. Rantzer, E. Davison, P. Young, and J. Doyle. A formula for computation of the real stability radius. *Automatica*, 31:879–890, 1995.

[17] M. Rostami. New algorithms for computing the real structured pseudospectral abscissa and the real stability radius of large and sparse matrices. *SIAM Journal on Scientific Computing*, 37:5447–5471, 2015.

[18] K. Schreiber. *Nonlinear eigenvalue problems: Newton-type methods and nonlinear Rayleigh functionals*. PhD thesis, TU Berlin, 2008.

[19] F. Tisseur and N. J. Higham. Structured pseudospectra for polynomial eigenvalue problems with applications. *SIAM Journal on Matrix Analysis and Applications*, 23(1):187–208, 2001.

[20] L.N. Trefethen and M. Embree. *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*. Princeton University Press, 2005.

[21] R. Van Beeumen, K. Meerbergen, and W. Michiels. Computing a partial Schur factorization of nonlinear eigenvalue problems using the infinite Arnoldi method. *SIAM Journal of Matrix Analysis and Applications*, 36(2):820–838, 2015.

[22] Z. Wu and W. Michiels. Reliably computing all characteristic roots of delay differential equations in a given right half plane. *Journal of Computational and Applied Mathematics*, 236:2499–2514, 2012.

Appendix

Proof of Theorem 4.1

The real case (i) is known [14]. So consider the case (ii), where λ_{RM} is genuinely complex. For smooth matrix valued functions $E_i(t)$, with $E_i(0) = E_i$ and $\|E_i(t)\|_2 \leq 1$ then, using Lemma 3.1, the classical derivative formula for a simple eigenvalue of the nonlinear eigenvalue problem with matrices $A_i - \frac{\epsilon}{w_i} E_i(t)$ leads to

$$\left.\frac{d\lambda}{dt}\right|_{t=0} = c\, x^* \left(\sum_{i=1}^{m} \frac{\epsilon}{w_i} p_i\left(\lambda_{\mathrm{RM}}\right) \dot{E}_i(0) \right) y \tag{7.1}$$

with c a positive constant. Now let $\dot{E}_i(t) \equiv 0$ for all $i = 1, \ldots, m$ except i_*. Formula (7.1) gives

$$\left.\frac{d\lambda}{dt}\right|_{t=0} = c\, \frac{\epsilon}{w_i} p_i\left(\lambda_{\mathrm{RM}}\right) x^* \dot{E}_{i_*}(0)\, y. \tag{7.2}$$

Recall that the x and y are the left and right eigenvectors, respectively, both of unit norm and scaled according to (4.1). Let $p_{i_*}(\lambda_{\mathrm{RM}}) = r\mathrm{e}^{\mathrm{i}\theta}$ (note that we have assumed $r \neq 0$); set $z = \mathrm{e}^{\mathrm{i}\theta} y$ (a rotated eigenvector) and denote by $x = x_R + \mathrm{i}x_I$, $y = y_R + \mathrm{i}y_I$ and $z = z_R + \mathrm{i}z_I$. Then we define the matrices X, Y (and Z coherently) as we did in Theorem 3.3.

By the rightmost property of λ_{RM}, (7.2) gives $0 \geq \Re\left(x^* \dot{E}_{i_*}(0)\, z\right)$, which implies

$$0 \geq x_R^T \dot{E}_{i_*}(0) z_R + x_I^T \dot{E}_{i_*}(0) z_I = \mathrm{trace}(X^T \dot{E}_{i_*}(0) Z) = \left\langle X\, Z^T, \dot{E}_{i_*}(0) \right\rangle. \tag{7.3}$$

Assume, by contradiction, that $\|E_{i_*}\|_2 = \|E_{i_*}(0)\|_2 < 1$; then we can always find $\dot{E}_{i_*}(0)$ such that violates (7.3) and increase the real part of the rightmost eigenvalue λ_{RM}. By continuity of the norm the matrix $E_{i_*}(t)$ would have spectral norm smaller than 1 in a sufficiently

small neighbourhood of $t = 0$ so it would determine an admissible perturbation. As a consequence, being our argument independent of i_* we have that

$$\|E_i\|_2 = 1 \qquad \text{for all } i = 1, \dots, m.$$

For the rank-2 equivalent extremizer property, let $\widehat{U}, \widehat{V} \in \mathbb{R}^{n\times n}$ be orthogonal matrices whose first two columns $U, V \in \mathbb{R}^{n\times 2}$ span the range of X, and of Y and Z respectively:

$$X = \widehat{U}\begin{pmatrix} S \\ 0 \end{pmatrix} = US, \quad Z = \widehat{V}\begin{pmatrix} T \\ 0 \end{pmatrix} = VT \quad \text{with } S, T \in \mathbb{R}^{2\times 2}.$$

Let $E_{i_*}(t)$, for small $t \geq 0$, be a continuously differentiable path, which we write as

$$E_{i_*}(t) = \widehat{U}\widehat{R}(t)\widehat{V}^T, \qquad \|\widehat{R}(t)\|_2 = 1$$

where $E_{i_*}(0) = E_i$.

Let $R(t) \in \mathbb{R}^{2\times 2}$ denotes the left upper 2×2 block of $\widehat{R}(t)$, and set

$$E_{i_*}(0) = UR(0)V^T := \tilde{E}_{i_*} \qquad \text{with} \quad \|\tilde{E}_{i_*}\|_2 \leq 1. \tag{7.4}$$

Observe that, as a consequence of (7.3),

$$0 \geq \operatorname{trace}(X^T \dot{E}_{i_*}(0) Z) = \operatorname{trace}(S^T \dot{R}(0) T). \tag{7.5}$$

This inequality holds for every path $\widehat{R}(t)$ of unit norm such that $\|R(0)\|_2 = 1$, otherwise we could choose $\dot{R}(0)$ arbitrarily; thus $\|\tilde{E}_{i_*}\|_2 = 1$. Let us write

$$R(0) = P\Sigma Q^T, \qquad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & \sigma_2 \end{pmatrix},$$

with orthogonal 2×2 matrices P, Q and $0 \leq \sigma_2 \leq 1$.

Assume $\sigma_2 < 1$; choosing the matrices $R_k(t) = P\Sigma_k(t)Q$ with

$$\Sigma_1(t) = \begin{pmatrix} 1 & 0 \\ 0 & \sigma_2 \pm t \end{pmatrix}, \quad \Sigma_2(t) = \begin{pmatrix} 1 - at^2 & \pm t \\ 0 & \sigma_2 \end{pmatrix}, \quad \Sigma_3(t) = \begin{pmatrix} 1 - at^2 & 0 \\ \pm t & \sigma_2 \end{pmatrix},$$

it is easily verified that if $\sigma_2 < 1$ it is possible to find $a > 0$ such that both singular values of the matrices $R_{2,3}(t)$ have spectral norm not larger than one in a sufficiently small neighborhood of $t = 0$.

Next we find that the inequality $0 \leq \operatorname{trace}(S^T \dot{R}(0) T)$ (see (7.5)) implies

$$0 = \operatorname{trace}(S^T P e_\ell e_j^T Q^T T) = e_j^T Q^T T S^T P e_\ell \quad \text{for } (\ell, j) \in \{(2,2), (1,2), (2,1)\},$$

where $e_1 = (1 \quad 0)^T$ and $e_2 = (0 \quad 1)^T$. As a consequence

$$Q^T T S^T P = \begin{pmatrix} \eta & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{with real } \eta$$

and therefore S or T must be of rank 1. This yields that x_R and x_I are linearly dependent or z_R and z_I are linearly dependent, an occurrence which can be discarded by the arguments in [3]; this implies $\sigma_2 = 1$.

Finally, $\sigma_1 = \sigma_2 = 1$ implies that $\widehat{R}_{21}(0) = 0$, since otherwise one of the first two columns of $\widehat{R}(0)$ has norm larger than one, in contradiction to $\|\widehat{R}(0)\|_2 = 1$. Hence

$$\widehat{R}(0) = \begin{pmatrix} R(0) & 0 \\ 0 & \widehat{R}_{22} \end{pmatrix},$$

with a matrix $\widehat{R}_{22}$ such that $\|\widehat{R}_{22}\|_2 \leq 1$.

Repeating this argument for all indexes $i = 1, \ldots, m$ yields that

$$\left(\sum_{i=1}^{m} \left(A_i + \widetilde{\delta A}_i \right) p_i(\lambda_{\mathrm{RM}}) \right) y = 0, \quad x^* \left(\sum_{i=1}^{m} \left(A_i + \widetilde{\delta A}_i \right) p_i(\lambda_{\mathrm{RM}}) \right) = 0$$

with $\widetilde{\delta A}_i = \dfrac{\epsilon}{w_i} \tilde{E}_i$ of rank-2 for all i, and the theorem is proved.

Proof of Theorem 4.2

We make use of the following result (see [3]).

Lemma 7.1. *(a) Let B be a real square matrix. Then we have* $\mathrm{trace}(BZ) = 0$ *for every skew-symmetric matrix Z if and only if B is symmetric.*

(b) Let B be a symmetric matrix. Then we have $\mathrm{trace}(BM) \geq 0$ *for every symmetric positive semidefinite matrix M if and only if B is positive semidefinite.*

Property (ii) has been obtained in the proof of Theorem 4.1; property (i) is a consequence of the representation of a rank-2 matrix with both singular values equal to 1 and image given by $\mathrm{span}(U)$ and kernel $\mathrm{span}(V^{\perp})$.

For the proof of (iii) we make use of Lemma 7.1 and similar arguments to those used in [3].

For a chosen $i_* \in \{1, \ldots, m\}$ consider the smooth matrix valued function

$$E_i(t) = U Q_{i_*} e^{tS} V^T$$

where S is a skew-symmetric matrix and let all the other matrices be unchanged. Denote by $\lambda(t)$ the eigenvalues $\lambda(t)$ with $\lambda(0) = \lambda_{\mathrm{RM}}$.

By local optimality, we have (see (7.3))

$$\begin{aligned} 0 \;&\geq\; \Re \left(p_i(\lambda_{\mathrm{RM}})\, x^* \dot{E}_{i_*}(0)\, y \right) = \mathrm{trace}(X^T \dot{E}_{i_*}(0) Y \Gamma_{i_*}) \\ &=\; \left\langle X\, \Gamma_i^T\, Y^T, U Q_{i_*} S V^T \right\rangle = \left\langle Q_{i_*}^T\, U^T\, X\, \Gamma_{i_*}^T\, Y^T V, S \right\rangle \end{aligned}$$

which holds for every skew-symmetric matrix S. By Lemma 7.1, $Q_{i_*}^T\, U^T\, X\, \Gamma_{i_*}^T\, Y^T V$ is symmetric.

We next consider the smooth matrix valued function

$$E_{i_*}(t) \;=\; U Q_{i_*} e^{-tM} V^T \quad \text{with a symmetric positive semi-definite matrix } M$$

where M is a symmetric positive semi-definite matrix and let all the other matrices be unchanged.

Then we obtain

$$0 \geq \Re(p_i(\lambda_{\mathrm{RM}})\, x^* \dot{E}_{i_*}(0) y) = -\left\langle Q_{i_*}^T\, U^T\, X\, \Gamma_{i_*}^T\, Y^T V, M \right\rangle$$

for every symmetric positive semi-definite matrix M. By Lemma 7.1 and invertibility $Q_{i_*}^T\, U^T\, X\, \Gamma_{i_*}^T\, Y^T V$ is positive-definite. Being the argument independent of i_* the proof is complete.